United States
Environmental
Protection Agency

Science Advisory
Board (A-101)

EPA-SAB-EHC-93-007
February 1993

AN SAB REPORT: SUPERFUND SITE HEALTH RISK ASSESSMENT GUIDELINES

REVIEW OF THE OFFICE OF SOLID WASTE AND EMERGENCY RESPONSES DRAFT RISK ASSESSMENT GUIDANCE FOR SUPERFUND HUMAN HEALTH EVALUATION MANUAL BY THE ENVIRONMENTAL HEALTH COMMITTEE

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

February 22, 1993

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

EPA-SAB-EHC-93-007

Honorable Carol M. Browner
Administrator
U.S. Environmental Protection Agency
401 M Street, S.W
Washington, DC 20460

Subject: Science Advisory Board Review of the Office of Solid Waste and Emergency Response draft *Risk Assessment Guidance for Superfund (RAGS), Human Health Evaluation Manual (HHEM),*

Dear Ms. Browner:

Early in the implementation of the Comprehensive Emergency Response Compensation and Liability Act, the EPA's Office of Solid Waste and Emergency Response (OSWER) decided to rely heavily on site-specific assessments of human and environmental risk to determine the need for remedial action, and to set protective cleanup levels. This approach was utilized because of the substantial differences in land use activities, terrain, hydrogeology, and nature and extent of contamination from site to site.

OSWER developed risk assessment guidance for Superfund to increase consistency in the way risk assessments are conducted within and across EPA Regional offices. The health risk assessment guidance for the Superfund program is codified in several documents, hereafter referred to collectively as the *Risk Assessment Guidance for Superfund, Human Health Evaluation Manual* (RAGS/HHEM). To ensure that the final RAGS/HHEM document reflects state-of-the-art technical guidance, and to comply with the recommendations of the Superfund 30-Day Study Task Force that OSWER should seek internal and external review of the Superfund risk assessment guidance, OSWER officials requested that the Science Advisory Board review selected issues addressed by the RAGS/HHEM document. Consequently, the Science Advisory

Board's Environmental Health Committee (EHC) met on April 7-8, 1992 in Bethesda, Maryland to review four broad issues relating to Superfund human health risk assessment:

a) Defining and calculating the Reasonable Maximum Exposure (RME) -- a conservative exposure case (i.e., well above the average) that is still within the range of possible exposures.

b) Assessing and dealing with exposures to multiple chemicals -- Using the Hazard Index (HI)/Hazard Quotient (HQ) to assess risk.

c) Reference doses in goal-setting -- use of chronic/sub-chronic RfDs and specific populations to set risk-driven remediation goals.

d) Short-term toxicity values -- use of appropriate defaults for characterizing less-than-lifetime exposure to toxicants.

The Committee found OSWER's attempts to improve the consistency of its risk assessment methodology praiseworthy. The OSWER has (understandably, given the range and complexity of the issues addresed) not succeeded fully in meeting their goals, but has made a good start. The Committee's findings note where a redirection in approach is called for, and provide advice where possible.

The Committee is of the opinion that there are some serious conceptual and practical problems with the proposal to calculate an RME based on an upper confidence limit (UCL) on the average concentration at a site. Given the proposed methods for computing the UCL, its underlying statistical assumptions, and problems in dealing with the spatial distribution of contamination *via-a-vis* the relative frequency with which people are likely to visit various parts of the site, the resulting UCL may have little, if any, relation to actual concentrations to which people may be exposed at a site. The Committee recommends that the EPA move to a distributional approach for calculating the RME, i.e., develop distributions for each of the terms or variables needed to calculate individual exposures and their distributions. These distributions determine a subsequent distribution for exposure, which can be calculated using Monte Carlo methods. A particular percentile of this exposure distribution, such as the 90th percentile, could then be used as the definition of the RME. To implement such an approach, EPA should develop default distributions for exposure parameters unlikely to vary

significantly from site to site. In situations where site-specific conditions may be unique, however, the collection of site-specific data is encouraged. Until this approach can be put into use, a modification of the current approach may have to be used. The Committee agrees with OSWER that, as long as some type of mean concentration is to be employed to estimate human exposure, an arithmetic mean is more appropriate than a geometric mean.

The Committee is concerned about the approach of using Reference Dose (RfD)-derived Hazard Quotients/Hazard Indices as a basis for adding "risks" from exposure to complex mixtures; it is not truly risks which are being added when the proposed approach is used. Quantitative applications using dose-response data (not the "point" data represented by Lowest Observed Adverse Effects Level/No Observed Adverse Effects Level (LOAEL/NOAEL)-derived RfDs) would be preferable, as would the use of alternatives to the current default approaches that assume risk additivity. The use of the HI itself can be misleading, and it should be used as a "fallback," with full recognition of its flaws, only when more refined toxicological data are not available.

The HHEM recommends using the RfD as the toxicity criterion for each of the other effects believed to be caused by a given agent in a chemical mixture. This proposal does not deal with effect interactions, nor with the fact that many RfDs are based on non-specific endpoints which can stem from many different causes. It is not nearly as seriously flawed as the alternatives presently in use, however, and despite its flaws is an improvement.

Three approaches for using RfDs to develop risk-based remediation goals for contaminated soil, involving differing exposure scenarios and target populations, were presented. The most supportable of these uses a 30-year time-weighted average with a chronic RfD; differences between the three approaches are not dramatic however, and OSWER should study all three approaches to verify its ultimate choice (or range of choices).

The Committee sees no particular problems in the existing approach for dealing with short-term toxicity estimates. OSWER should take cognizance of the EPA-sponsored work at the National Academy of Sciences on Community Emergency Exposure Levels, and of the work on Emergency Response Planning Guidelines by the American Conference of Governmental Industrial Hygienists.

We appreciate having been given the opportunity to address these issues, and look forward to receiving your response to our comments.

Sincerely,

Raymond C. Loehr

Dr. Raymond Loehr, Chair
Science Advisory Board

Arthur C. Upton

Dr. Arthur C. Upton, Chair
Environmental Health Committee

ABSTRACT

The Office of Solid Waste and Emergency Response (OSWER) developed the *Risk Assessment Guidance for Superfund (RAGS), Human Health Evaluation Manual (HHEM), Part A--Baseline Risk Assessment* (December 1989), supplemented in March 1991 with *Standard Default Exposure Factors* guidance, and *Part B--Development of Risk-based Preliminary Remediation Goals*, and *Part C--Risk Evaluation of Remedial Alternatives* in December 1991 (as interim documents) to guide Agency staff performing site-specific assessments of human and environmental risk to determine the need for remedial action. At the request of OSWER, the Science Advisory Board's Environmental Health Committee (EHC) met on April 7-8, 1992 to review four broad issue areas relating to Superfund human health risk assessment: a) Defining and calculating the Reasonable Maximum Exposure (RME); b) Assessing and dealing with exposures to multiple chemicals and using the Hazard Index (HI)/Hazard Quotient (HQ) to assess risk; c) reference doses in remediation goal-setting; and d) Use of appropriate defaults for characterizing less-than-lifetime exposure to toxicants. The Committee found OSWER's attempts to improve the consistency of its risk assessment methodology to be praiseworthy and a good start, but noted areas where a revised approach is recommended.

The Committee is of the opinion that there are some serious conceptual and practical problems with the proposal to calculate an RME based on an upper confidence limit (UCL) on the average concentration at a site. The EHC recommends that the EPA move to a distributional approach to calculating the RME, i.e., developing distributions for each of the terms or variables needed to calculate individual exposures and their distributions. Given the difficulty in interpreting the RME as presently calculated, the Committee recommends that some type of 'most reasonable' estimate of exposure also be calculated and made available to risk managers along with the RME. The Committee agrees with OSWER that, as long as some type of mean concentration is to be employed to estimate human exposure, an arithmetic mean is more appropriate than a geometric mean.

The Committee is concerned about the approach of using RfD-derived Hazard Quotients/Hazard Indices as a basis for adding "risks" from exposure to complex mixtures. Quantitative applications using dose-response data (not the "point" data represented by LOAEL/NOAEL-derived RfDs) would be preferable, as would the use of alternatives to the current default approaches that assume risk additivity. The use of the HI itself can be misleading, and it should be used as a "fallback," with full recognition of its possible inapplicability, only when more refined toxicological data are not available. Interpretation of an HI greater than 1 can vary depending on several toxicological factors. Although it is likely that risk increases as the HI exceeds 1, we can not state (without a more complete understanding of interaction mechanisms) how rapidly this increase occurs, nor can we rely on HI-based comparisons of risks when the HIs are greater than 1.

Three approaches for using RfDs to develop risk-based remediation goals for contaminated soil were presented. The most supportable of these uses a 30-year time-weighted average with a chronic RfD; differences between the three approaches are not dramatic however, and OSWER should study all three approaches to verify its ultimate choice (or range of choices).

The Committee sees no particular problems in the existing approach for dealing with short-term toxicity estimates. OSWER should take cognizance of the EPA-sponsored work at the National Academy of Science on Community Emergency Exposure Levels, and of the work on Emergency Response Planning Guidelines by the American Conference of Governmental Industrial Hygienists.

KEYWORDS: complex mixtures; exposure; Hazard Index (HI); Hazard Quotient (HQ); reasonable maximum exposure; risk assessment; Reference Dose (RfD); Superfund; site assessment

NOTICE

This report has been written as a part of the activities of the Science Advisory Board, a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The Board is structured to provide balanced, expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

ENVIRONMENTAL PROTECTION AGENCY
SCIENCE ADVISORY BOARD
ENVIRONMENTAL HEALTH COMMITTEE

April 7-8, 1992

CHAIRMAN
Dr. Arthur C. Upton, New York University Institute of Environmental Medicine, New York City, NY

MEMBERS & CONSULTANTS
Dr. William Bunn, Mobil Oil, Princeton, NJ

Dr. Kenny Crump, Clement International Corporation, Ruston, LA

Dr. Paul Deisler, Retired, Shell Oil, Houston, TX

Dr. Adam Finkel, Resources for The Future, Washington, DC

Dr. David Gaylor, National Center for Toxicological Research, Jefferson, AR

Dr. Rolf Hartung, University of Michigan, Ann Arbor, MI

Dr. Marshall Johnson, Jefferson Medical College, Philadephia, PA

Dr. Nancy K. Kim, New York Department of Health, Albany, NY

Dr. Rogene Henderson, Lovelace Inhalation Toxicology Research Institute, Albuquerque, NM

Dr. Richard Monson, Harvard University School of Public Health, Boston, MA

Dr. Martha J. Radike, University of Cincinnati Medical Center, Cincinnati, OH

Dr. Bernard Weiss, University of Rochester School of Medicine, Rochester, NY

Dr. Ronald Wyzga, Electric Power Research Institute, Palo Alto, CA

DESIGNATED FEDERAL OFFICIAL
Mr. Samuel Rondberg, Environmental Health Committee, Science Advisory Board, U.S. Environmental Protection Agency

STAFF SECRETARY
Ms. Mary L. Winston, Science Advisory Board, U.S. Environmental Protection Agency

ASSISTANT STAFF DIRECTOR
Mr. Robert Flaak, Science Advisory Board, U.S. Environmental Protection Agency

STAFF DIRECTOR
Dr. Donald G. Barnes, Science Advisory Board, U.S. Environmental Protection Agency

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

Early in the implementation of the Comprehensive Emergency Response Compensation and Liability Act (CERCLA), the EPA's Office of Solid Waste and Emergency Response (OSWER) decided to rely heavily on site-specific assessments of human and environmental risk to determine the need for remedial action. OSWER developed the *Risk Assessment Guidance for Superfund (RAGS), Human Health Evaluation Manual (HHEM), Part A--Baseline Risk Assessment* (December 1989), supplemented in March 1991 with *Standard Default Exposure Factors* guidance, and *Part B--Development of Risk-based Preliminary Remediation Goals*, and *Part C--Risk Evaluation of Remedial Alternatives* in December 1991 (as interim documents).

At the request of OSWER program officials, the Science Advisory Board's Environmental Health Committee met on April 7-8, 1992 to review four broad issue areas relating to Superfund human health risk assessment:

a. Defining and calculating the Reasonable Maximum Exposure (RME) -- The goal of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

b. Assessing and dealing with exposures to multiple chemicals -- Using the Hazard Index (HI)/Hazard Quotient (HQ) to assess risk.

c. Reference doses in goal-setting -- use of chronic/sub-chronic Reference Doses (RfDs) and specific populations to set risk-driven remediation goals.

d. Short-term toxicity values -- use of appropriate defaults for characterizing less-than-lifetime exposure to toxicants.

The Committee found OSWER's attempts to improve the consistency of its risk assessment methodology praiseworthy. The OSWER has (understandably, given the range and complexity of the issues addressed) not succeeded fully in meeting their goals, but has made a good start. The Committee's major findings follow, specifically noting where a redirection in approach is called for, and providing advice where possible.

The Committee is of the opinion that there are some serious conceptual and practical problems with the proposal to calculate an RME based on an upper confidence limit (UCL) on the average concentration at a site. The current approach assumes that the samples taken are representative of those areas where exposures are most likely to occur. The RME is also a function of the number of samples available; a larger number of samples will result in a smaller RME even if the samples are not representative of exposure opportunities. This approach does not deal with "hot spots" at a site which could cause visitors/residents to be exposed to levels significantly higher than the UCL. Given the proposed methods for computing the UCL, its underlying statistical assumptions, and problems in dealing with the spatial distribution of contamination *via-a-vis* the relative frequency with which people are likely to visit various parts of the site, the resulting UCL may have little, if any, relation to actual concentrations to which people may be exposed at a site. The Committee recommends that the EPA move to a distributional approach to calculating the RME, i.e., developing distributions for each of the terms or variables needed to calculate individual exposures and their distributions. These distributions determine a distribution for exposure, which can be calculated using Monte Carlo methods. A particular percentile of this exposure distribution, such as the 90th percentile, could be used as the definition of the RME. Kriging and triangulation are two statistical methods for quantifying spatial distributions of contaminant concentrations which could be used as part of this approach to address the issue of hot spots.

To implement a distributional approach, EPA should develop default distributions for exposure parameters unlikely to vary significantly from site to site. The collection of site-specific data is encouraged in instances where site-specific conditions may be unique, however. Until this approach can be put into use, a modification of the current approach may have to be used. Given the difficulty in interpreting the RME as presently calculated, the Committee recommends that some type of 'most reasonable' estimate of exposure also be calculated and made available to risk managers along with the RME. The Committee agrees with OSWER that, as long as some type of mean concentration is to be employed to estimate human exposure, an arithmetic mean is more appropriate than a geometric mean.

The Committee is concerned about the approach of using RfD-derived Hazard Quotients/Hazard Indices as a basis for adding "risks" from exposure to complex mixtures; it is not truly risks which are being added when the proposed approach is used. Quantitative applications using dose-response data (not the "point" data

represented by LOAEL/NOAEL-derived RfDs) would be preferable, as would the use of alternatives to the current default approaches that assume risk additivity.

The use of the HI itself can be misleading, and it should be used as a "fallback," with full recognition of its possible inapplicability, only when more refined toxicological data are not available. The condition "HI = 1" defines a "threshold of concern" that is not shared by any other value for HI, and for which, under specified conditions, the uncertainty in the HI approach is no greater than that of the component RfDs. The HI approach is invalid, however, if the chemicals in the mixture cannot be fully characterized by a combination of dilution-type interactions and independent mechanisms of action. Interpretation of an HI greater than 1 can vary depending on several toxicological factors. Although it is likely that risk increases as the HI exceeds 1, we can not state (without a more complete understanding of interaction mechanisms) how rapidly this increase occurs, nor can we rely on HI-based comparisons of risks when the HIs are greater than 1.

The HHEM recommends using the RfD as the toxicity criterion for each of the other effects believed to be caused by a given agent in a chemical mixture. This proposal does not deal with effect interactions, nor with the fact that many RfDs are based on non-specific endpoints which can stem from many different causes. It is not nearly as seriously flawed as the alternatives presently in use, however, and despite its flaws is an improvement.

Three approaches for using RfDs to develop risk-based remediation goals for contaminated soil, involving differing exposure scenarios and target populations were presented. The most supportable of these uses a 30-year time-weighted average with a chronic RfD; differences between the three approaches are not dramatic however and OSWER should study all three approaches to verify its ultimate choice (or range of choices).

The Committee sees no particular problems in the existing approach for dealing with short-term toxicity estimates. OSWER should take cognizance of the EPA-sponsored work at the National Academy of Science on Community Emergency Exposure Levels, and of the work on Emergency Response Planning Guidelines by the American Conference of Governmental Industrial Hygienists. The method proposed by Region 6 for setting short-term air action levels calls for the possible use of OSHA standards, such as Permissible Exposure Limits (PELs) and Short-Term Exposure

Levels (STELs). The data used to derive any OSHA standard so used should be examined to see if the same data can be used to derive an appropriate RfD. The use of EPA derivation methods would help promote consistency across various hazardous substances; hence the use of EPA methods should be encouraged. Where this is not possible or practical, the use of health data on which OSHA standards have been based could be considered, taking into account differences between the worker population and the general population.

2. INTRODUCTION

2.1 Background

Risk assessment is an essential component of the Superfund site remediation process. Early in the implementation of the Comprehensive Emergency Response Compensation and Liability Act (CERCLA), the EPA's Office of Solid Waste and Emergency Response (OSWER) decided to rely heavily on site-specific assessments of human and environmental risk to determine the need for remedial action, to identify contaminants of concern and critical exposure pathways, and to determine protective cleanup levels. This approach was utilized because of the substantial differences in land use activities, terrain, hydrogeology, and nature and extent of contamination from site to site. OSWER believed that decisions regarding the protectiveness of contaminant concentrations in the environment were best made on the basis of specific site circumstances.

OSWER developed risk assessment guidance for Superfund to reflect extensive experience obtained from conducting health and environmental risk assessments at Superfund sites, utilizing existing Agency risk assessment methods and data bases. The guidance is designed to increase consistency in the way risk assessments are conducted within and across EPA Regional offices. The health risk assessment guidance for the Superfund program is presented in several documents. *Risk Assessment Guidance for Superfund (RAGS), Human Health Evaluation Manual (HHEM), Part A--Baseline Risk Assessment*, was published as interim final in December 1989. This Part was supplemented in March 1991 with *Standard Default Exposure Factors* guidance (OSWER Directive 9285.6-03). Two additional Parts of the Human Health Evaluation Manual strengthen the relationship among human health evaluation, cleanup goals, and remedy selection. *Part B--Development of Risk-based Preliminary Remediation Goals*, and *Part C--Risk Evaluation of Remedial Alternatives* were published as interim documents in December 1991. In the near future, all three Parts of HHEM will be integrated into a single final HHEM document that will incorporate new policy information and other technical guidance that has been issued by the Agency since 1989, as well as comments received from field users.

Unlike risk assessments conducted by other EPA Program Offices that often focus on single contaminants or single exposure pathways, Superfund assessments must address multiple contaminants and multiple pathways for each site on the National Priorities List (currently more than 1200 sites). The Superfund program deals

with diverse problems including waste from spills, illegal dumping, landfills, surface impoundments, drum/container storage, as well as other contaminant sources. The Superfund program typically conducts risk assessments at well over 100 sites per year, in all regions of the country, under varied environmental conditions, and for multiple land uses. There is strong public and political pressure on this program to address these sites as quickly as possible. Efforts to characterize baseline health risks at a site typically take two to three months once data are in hand, but are sometimes delayed because of the need to collect better sampling data, or negotiations with "potentially responsible parties" over land use, exposure assumptions, and chemical toxicity.

To address the needs of Superfund risk assessors, Superfund's risk assessment guidance must be flexible enough to encompass the wide variety of conditions present at sites in all Regions yet specific enough to assure a reasonable degree of consistency. Assessors must balance the pressures to gather additional data (usually resource intensive) with the need to move quickly toward site cleanup. Risk assessors and site managers are often required to make real-time cleanup decisions at Superfund sites with imperfect data.

In the interest of continuous improvement, to ensure that the final RAGs-HHEM document reflects state-of-the-art technical guidance, and to comply with the recommendations of the Superfund 30-Day Study Task Force that OSWER should seek internal and external review of the Superfund risk assessment guidance, program officials requested that the Science Advisory Board review selected issues addressed by the Risk Assessment Guidance documents noted above.

2.2. Charge To The Committee

Four broad issue areas relating to Superfund human health risk assessments were identified for review; within each issue, specific questions were posed to the Committee.

a) Reasonable Maximum Exposure: Superfund's approach for calculating the Reasonable Maximum Exposure (RME) is presented in HHEM Part A (Chapters 6 and 8) and in its supplement "Standard Default Exposure Factors" (OSWER Directive 9285.6-03). The goal of the RME is to

estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

1) The HHEM Part A identifies the arithmetic mean concentration, within an appropriate averaging zone, as the measure of concentration that generally represents the integrated long-term exposure that will be received by an individual within that zone. Is this approach reasonable?

2) Part A also indicates that given the typical distribution and spatial/temporal variability of contaminant concentration and the often limited number of samples collected at Superfund sites, it is important that the health risk assessments explicitly address uncertainty in the mean concentration. Do 95% confidence limits on the mean concentration provide an appropriate tool for addressing uncertainty in concentration?

3) Superfund is currently investigating alternate approaches for characterizing contaminant concentrations, such as kriging and triangulation methods. Are these alternative approaches for estimating average concentrations (or others) appropriate to consider for future guidance?

4) Superfund guidance indicates that valid site-specific information on exposure factors -- particularly human behavior patterns -- be used in exposure assessments. In the absence of site-specific survey data, or in cases where the assessment must determine projected changes in land use, guidance has relied on survey data for other populations, commonly at the national level. Is this approach reasonable?

5) The Agency's new "Guidelines for Exposure Assessment" (Section 5.3.5.1) discuss three approaches towards making estimates of exposure for highly exposed individuals. First, upper percentiles of population exposure may be directly determined from surveys where direct measurement data on exposure were obtained. Second, mathematical simulation techniques can be used to

combine adequate distributional data of individual exposure factors to estimate the overall distribution of exposure. Third, an estimate may be constructed using a combination of upper-bound and mid-range values for exposure factors. In this case, the few most variable factors would generally be set to upper-range values and the remaining factors set to mid-range values. Experience indicates that under appropriate circumstances each of the approaches can provide useful information on Superfund site risk. However, since data on exposed populations are limited, HHEM Part A provides additional details for the third tier situation. Is the Superfund approach of combining high-end and mid-range values presented in HHEM Part A (Chapter 6, Sections 6.1.2 and 6.4) consistent with the third approach for estimating high-end exposures mentioned above and described in the "Guidelines for Exposure Assessment" (Section 5.3.5.1)?

b) Exposures to Multiple Chemicals: The current approach to assessment of hazard or risk of concurrent exposure to multiple chemicals (HHEM Part A, Chapter 8), is based on the assumption of dose additivity, as recommended in EPA's "Guidelines for Health Risk Assessment of Chemical Mixtures." For carcinogens, simple additivity of risk is used. For non-carcinogens, a Hazard Index (HI) approach has been developed. In applying the HI approach, the potential for non-carcinogenic effects is evaluated by comparing estimated exposure to a reference dose (RfD). The resulting ratio for each chemical is called a hazard quotient. If the sum of these hazard quotients (called a hazard index or HI) for several chemicals with the same toxic endpoint exceeds unity, there is a concern for potential adverse effects.

6) Given what is currently understood about the potential interactive effects of chemicals, is it appropriate to add risk estimates for multiple contaminant exposures (i.e., calculation of Hazard Indices for chemicals with similar toxic endpoints, and simple additivity of cancer risks)?

7) Given the uncertainties in determining HIs, is it appropriate to use HI values greater than a specific and constant number (i.e., "1") as a threshold of concern?

8) What does an HI greater than 1 represent when the hazard quotients used to calculate the HI are individually less than 1 (i.e., when the estimated "dose" from an individual chemical does not exceed the RfD)?

9) Existing guidance specifies that chemicals should be grouped according to their major toxic effects, including those seen at doses or exposures higher than those associated with the critical effect (upon which the RfD was based). As a conservative and simplifying step, the guidance recommends that the RfD be used as the toxicity criterion for each of the other effects believed to be caused by that chemical. Is this a reasonable approach? The issue paper provided to the Committee (*Derivation of Effect-Specific RfDs and Their Use in Risk Assessment for Chemical Mixtures*) presents an alternative approach. Is this approach (or are others) reasonable to consider for future guidance?

c) Reference Doses in Goal-setting: During the development of the HHEM Part B, two approaches were considered for using RfDs in setting risk-based remediation goals in soil: 1) comparison of a 6-year, childhood exposure to contaminants in soil with a sub-chronic RfD; and, 2) comparison of a 30-year, time-weighted average exposure to contaminants in soil (including exposures to both children and adults) with a chronic RfD. At the time, the second approach was chosen, as it provided a more conservative cleanup goal. Since that time, a third approach has been proposed: comparison of a 6-year, childhood exposure with a chronic RfD.

10) Given EPA's definition of an RfD [i.e., "an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime"], what is the panel's opinion regarding the

appropriateness of the second approach? Under what conditions should other approaches be considered?

d) Short-term Toxicity Values: Current guidance in HHEM Part C (Appendix C, p. 49) refers risk assessors to the Superfund Technical Support Center at ECAO-Cincinnati to obtain any toxicity criteria needed for risk assessment of less-than-lifetime exposures. Since data on assessing short-term exposure risks are extremely limited and no Agency-wide guidance is available, ECAO often must derive interim toxicity criteria based on the methods outlined in the *Interim Methods for Development of Inhalation Reference Concentrations* (p. 4-37).

11) In the absence of chemical- and duration-specific data, is this method reasonable? Are alternate procedures available? -

There is currently no national guidance for setting short term air action levels that would guide activity (and emergency shutdown) during remedial action at Superfund sites. EPA Regions are currently using various approaches for deriving such levels. For example, Region 6 has issued a policy statement that discusses the derivation of air action levels from EPA's chronic health risk values; from OSHA Permissible Exposure Limits, Short Term Exposure Levels, or Ceiling Values; or from RfD/RfC values.

12) Is the Region 6 approach an appropriate method for deriving short-term air action levels? Is there a more appropriate method for developing short-term air action levels that could be used to trigger shutdown of cleanup operations and/or the evacuation of the general population near a Superfund site should unanticipated releases occur (e..g., toxicity values relevant to 15-minute human exposures)?

In addition to the specific questions noted above, the OSWER requested that this review take into consideration the context within which these issues will be dealt with operationally, i.e., the pressures and real-time decisions that must be made at Superfund sites.

3. SPECIFIC FINDINGS

3.1 Reasonable Maximum Exposure–Issue One

3.1.1 Arithmetic Mean Concentration

It is difficult for the Committee to evaluate whether OSWER's use of the arithmetic mean concentration is appropriate, because OSWER has neither adequately explained the rationale for moving from an upper-bound estimate of concentration to a mean estimate (and the ramifications thereof), nor sufficiently documented over what measurements the mean is to be estimated. The following briefly discusses each of these two issues in turn.

a) The "appropriate" mean

OSWER is attempting to account for several fundamentally different kinds of uncertainty or error with one rather vague policy. The combination of "reasonable high-end" values for those parameters that vary across individuals (e.g., breathing rate, ingestion rates, body weight, contact rate, exposure frequency, exposure duration), with a mean concentration within an "appropriate averaging zone" co-mingles different kinds of variability (inter-individual versus spatial/temporal) with various sources of uncertainty into a single measure that is supposed to be "conservative" but not wildly so. This is a worthwhile end, but the means to that end are arbitrary and unverified, and this measure has no consistent interpretation. First, OSWER needs to show (preferably by Monte Carlo simulations using both actual and stylized data sets) that the short-hand combination of "high-end" default values and mean concentration does in fact yield a reasonably conservative output[1]. More importantly, however, OSWER must demonstrate that the current spatial average (which needs to be defined) is in fact the long-term average for the "average individual" (or whatever individual it intends to model).[2]

---

[1] Such a demonstration should at least attempt to model the likely correlations among the behavioral variables (both positive and negative) and the possible correlations between the behavioral and concentration inputs.

[2] The burden of proof or of validation is definitely on EPA in this instance. Intuition tells us that the current spatial average itself is likely to change over time; to this, EPA adds the additional assumption that spatial averages and inter-individual lifetime averages are related

Clearly, part of this demonstration will come about from a better definition of "appropriate averaging zone." In addition, OSWER must also investigate whether even an "appropriate" spatial average yields a representative long-term concentration estimate for the relevant person or group. In other words, in addition to needing to specify over what area a person or group's "random walk" is occurring, OSWER needs to justify that the "walk" is in fact random. This might be more or less true if the issue was sporadic contact with an industrial site, but it may be inappropriate to average over a broad "averaging zone" in the case of widespread contamination in a residential area. Here the only "random walk" would be over small portions of each homeowner's property, not over an entire neighborhood.

Consequently, we do not believe that the proposed methods are consistent with *the Guidelines for Exposure Assessment*, particularly the proposed approach for estimating the appropriate concentration term in the RME. Using the mean, even allowing for its upper 95% confidence interval, does not estimate the reasonable maximum concentration to which an individual is exposed, but merely gives a confidence limit on the "average" site-wide concentration over the time period during which measurements were taken--if one assumes that the sites sampled were representative of the site of concern. If the sites sampled are not random, or are not part of a systematic design to characterize the site as a whole, we do not know how to interpret either the resulting average concentration or its upper 95% confidence interval, and we have no idea how it relates to the appropriate RME concentration.

One other technical question also needs to be addressed. It relates to the limitations of Haber's Rule[3], which forms the conceptual basis for time-weighted averaging. An implication of Haber's Rule is that the total amount of a dose governs the effect, and that the time-course of administration is irrelevant. Haber's Rule tends to work over relatively narrow ranges of exposures and duration, but, as the ranges increase, it begins to fail more and more. Because it does not take into account pharmacokinetics and biotransformations, its applications are limited, and the Agency should consider these limitations carefully whenever it proposes the use of time-weighted averages in its risk assessment procedures.

[3] Habers Rule: W= K x C x T, where W= Wirkung (a constant effect); K= a proportionality constant; C= concentration or dose/unit time; and T= duration of exposure.

To summarize, OSWER may have modeled the randomness of exposure to hazardous waste sites incorrectly, and could probably do a much better job by at least explaining whose behavior and exposures are being modeled, and why. The above issues notwithstanding, if EPA decides to use some type of mean concentration to estimate human exposure, the Committee agrees with OSWER that an arithmetic mean is more appropriate than other types of means, such as a geometric mean. The reasons for this are two-fold. First, an arithmetic mean may be more likely to correlate with health effects than a geometric mean. For example, if a person is exposed to 100 ppm of a chemical for 12 hours of the day and to 0.001 ppm for the other 12 hours per day, the arithmetic mean exposure is 50 ppm and the geometric mean exposure is 0.316 ppm. The arithmetic mean clearly reflects higher exposure levels that most likely have a greater influence on any health response. The second reason is that the geometric mean of a group of samples is highly sensitive to values assigned to analytical results below the detectable level of the assay. Theoretically, if a single value is truly zero, then the geometric mean is likewise zero, regardless of the remaining values.

b) Implications of using the mean estimate

Under certain conditions, involving both statistical factors and science-policy judgments, the mean is the preferred estimator of an uncertain and/or variable quantity. Within the context of hazardous waste-site risk assessment, the mean would be appropriate if: (1) the judgment were made that population risk, rather than individual risk, were the relevant metric with which to measure hazard; or (2) the assessor had reason to believe that the (spatial and temporal) variability of the concentration term were sufficiently small so that upper-bound individual risks would not deviate significantly from the mean. In addition, the assessor would have to believe (3) that the uncertainty of the concentration term (due to measurement or modeling errors) was sufficiently well understood so that the mean was neither an overly conservative estimate of the representative concentration (i.e., not unduly affected by a few outliers) nor biased low due to small sample sizes, incorrect model form, etc. OSWER has attempted to address this third condition by using the upper confidence limit (UCL) of the mean, but the Committee does not regard this as an adequate solution (see discussion following concerning the use of a designated percentile on the distribution in place of a confidence limit of the mean).

In addition, OSWER needs to communicate better to senior managers and the public that adopting any type of mean concentration represents a fundamental shift in the Superfund program, making it more like the pesticides program, for example, and less like the air toxics program (which considers the risk to the "maximally exposed individual" to be of equal or greater import than the population average risk or the total number of "lives saved"). If OSWER believes that by "appropriate" definition of the "averaging zone" (see below), it is circumscribing the risk analysis to address a relatively "tight" concentration distribution that accounts for "reasonable high-end exposures," it needs to make that assumption explicit and provide some empirical support for it. Otherwise, it should acknowledge that the use of the mean exposure implicitly embodies the controversial, value judgment that the distribution of individual risks is irrelevant for cleanup decisions or comparative risk analysis.

3.1.2 Use of the 95% Confidence Limit/Estimating Exposure--Issues Two and Five

The HHEM addresses the estimation of the Reasonable Maximum Exposure (RME) as an alternative to use of the Maximally Exposed Individual (MEI). We applaud the consideration of this concept and believe it could be of considerable value in evaluating health risks at Superfund sites.

The proposed estimate of the RME in the HHEM is a product of several terms, including the upper 95% confidence interval on the arithmetic mean of the contaminant concentration at the site of concern. The mean is averaged over temporal and spatial scales and can be applied to concentrations in all media. Question two of the Charge asks if this 95% upper confidence limit is appropriate, question five, which is closely related, asks about the consistency of this estimate, and the approaches to combining high- and mid-range exposure values. Both issues are discussed below.

a) Use of the UCL on the mean concentration

If the sampling design allows an estimate of the average concentration, we do not believe that the upper 95% confidence interval of this average is appropriate, unless there is evidence that individual exposures are equally likely across all parts of the site over time periods similar to that over which the samples were obtained. Another problem with the current approach is that the

number of samples taken will have undue influence on the upper confidence limit even if a large number of samples will not lead to good estimates of actual exposure. For example, consider a site where contaminant concentrations are highly non-uniform: there may be a hotspot where concentrations are high, and the remainder of the site is not highly contaminated. Assume that the hotspot is highly attractive to visitors (e.g., a waterhole or pond). If there are many samples taken across other areas of the site, as well as the hotspot (n is large) then the mean is not a good estimate of exposure, and the upper confidence limit on the mean is a poor estimate of the RME despite the fact that there are many sample observations. Estimation of the RME cannot ignore the distribution of contaminants at the site and the distribution of individual behaviors which lead to exposure.

The preceding discussion implies that the estimate of an overall mean concentration at the site, or an upper confidence limit thereon as proposed in the RAGs draft document, can be inadequate for calculation of an RME. Rather, the spatial distribution of the concentration over the site must be considered along with a distribution reflecting the relative frequency with which people are likely to visit different parts of the site. For this reason, any summary measure of concentration (such as the average proposed in the RAGs document) that does not take into account the spatial distribution of the underlying samples is likely to be inadequate. Therefore, the Committee believes that the Agency should give strong consideration to incorporating methods, such as kriging or triangulation (procedures that are discussed in more detail in Section 3.1.3) that take into account the spatial distribution of contamination to characterize exposure.

A related problem with calculating an upper confidence limit for a mean, and one that can have very severe practical consequences, is the fact that the statistical assumptions required to calculate an UCL for a mean exposure are typically not met by sampling plans at a Superfund site. First, the statistical procedures generally require that sample locations be selected randomly. However, in typical cases, sampling is done in several stages and is generally non-random. An initial screening may be performed to identify areas of particular concern, and follow-up sampling may be performed to characterize more completely the extent of contamination. Random sampling may not be used (or even be appropriate) at any stage of sampling.

Second, in order to calculate a statistical confidence limit for the mean, a particular distributional form must be assumed for the data (e.g., normal, log-normal, or Weibull). There is *a priori* no reason that sampling data should have a particular distributional form. Different distributional assumptions can sometimes produce quite disparate results.

In some cases, routine application of a standard procedure for calculating an UCL on the mean concentration can produce an anomalous result in which the UCL is literally thousands of times larger than the maximum observed value (MOV). In fact, such extreme behavior can even occur when obtaining a point estimate of the mean. When this anomaly occurs, it is likely to be due to the fact than one or more of the underlying statistical assumptions are violated, in which case the relationship between the computed UCL and the actual concentrations at the site are likely to be purely coincidental.

The RAGs manual suggests that the upper limit be replaced by the MOV in such cases. However, this default could result in a non-conservative estimate of the average concentration. Whenever the estimated mean, or UCL on the mean, exceeds the MOV, EPA should have concern as to whether or not the concentrations at the site have been adequately characterized.

A related problem that may be critical when calculating the mean concentration, or a UCL for the mean, is the statistical treatment of "non-detects" (samples in which no contamination is detected). These cases are often treated in an *ad hoc* fashion by assigning either the smallest concentration that could have been detected (the "detection limit") or some fixed fraction of the detection limit to these samples. However, the value assumed for non-detects can sometimes have an enormous effect upon the UCL for the mean concentration. However, likelihood methods are available for such data that do not require assigning a particular value to non-detects (see, for example, Crump, 1978). The Committee recommends that such methods be considered by EPA when calculating UCLs for mean concentrations from data containing non-detects.

Considering the estimation issue *de novo* by media of potential exposure, the proposed methodology is probably most relevant for soil contamination and dermal exposure. The calculation of the RME must consider the likely

exposure patterns of individuals visiting a site, and all parts of the site may not be equally likely of receiving a visit. An average concentration weighted by the likelihood of that sampling site being visited is a much better estimate of the concentration to be used for calculating the RME. This can be improved even further by looking at a distribution of average concentrations to which visitors at a site are exposed and taking a given percentile, such as the 90th percentile, of that distribution. We would advocate such an approach in place of the approach proposed in the HHEM. In addition, we believe that the alternative approach proposed above is more consistent with the exposure assessment guidelines, and is in the spirit of the *Exposure Factors Guidelines*.

b) Approaches to combining various exposure inputs

EPA is treating different factors in the exposure calculation in different manners (upper 95th confidence limit for one factor, best estimates for others, and some not clearly defined type of conservative limit for others), so that the result is very difficult to interpret. The Committee recommends that EPA move towards a full distributional approach in which distributions are developed for each of the terms in the exposure equation and a Monte Carlo analysis be applied to obtain the resulting distribution for exposure (and thus any desired percentile of this distribution, such as the 90th percentile). In this manner, EPA can consciously choose the desired degree of "conservatism." An EPA-sponsored effort in which such distributions are developed and applied to a few sites would illustrate the methods, expose the strengths and weaknesses of the methods in more detail, and provide guidance on the appropriate data needs to facilitate the calculation of the RME. To facilitate this, EPA should develop default distributions for terms that are not likely to vary greatly from site to site (amount of water consumed, body weight, etc.). It has been argued that the data are not always available to derive the estimates of the type we suggest. We believe that the costs of generating these data are not very great, and the burden of providing these data could be placed upon the Potentially Responsible Parties (PRPs). It would generally be in their interest to facilitate the estimation of the RME in place of the more conservative MEI.

For other media of exposure (air and water), the proposed methods need to be adapted because the RMEs may be associated less with exposures on

site than with exposures in areas adjacent to the site. The approach proposed in the HHEM ignores this possibility.

Given the difficulties in interpreting any single number as a measure of exposure, and in particular the RME as presently calculated, the Committee believes that some type of 'best' or 'most reasonable' estimate of exposure would provide useful additional information to a risk manager, and should be presented along with the "conservative" estimate we advocate above. If a Monte Carlo analysis is available, such an estimate could be obtained from the median of the uncertainty distribution. If a Monte Carlo approach has not been implemented, a 'best' or 'most reasonable' estimate could be developed by assigning 'best' or 'reasonable' values to each component of the estimating expression.

3.1.3 Characterizing Contaminant Concentrations--Issue Three

It is worth investigating alternate approaches for characterizing underground concentrations for possible use in future guidance. Of the two mentioned in the HHEM, kriging and triangulation, the former has been in use for many years, first in the extractive industries to map ore or coal deposits, for example, and more recently in characterizing underground contamination. The calculational method is well developed and the interpretation of the results is well understood. Typically, kriging yields not only concentration contours but also standard error contours, which facilitate developing sampling designs from a few initial samples, if these are available, as well as the determination of where additional samples may be needed to increase the confidence of definition of "hot spots" or other features of underground concentrations of contaminants. The Committee recommends that this technique be included in future guidance as a useful, already developed tool. Triangulation is a newer technique; different in its approach to data analysis from kriging, it nonetheless produces similar results and is worth further examination as a possible tool. Case studies, which illustrate the differences in results and their implications, should be developed before final judgements are made

3.1.4 Determining Exposure Factors--Issue Four

Although a well done survey is the best way to characterize behavior patterns of individuals working, living or otherwise present in the vicinity of Superfund sites, it

may not be necessary, in many cases, to carry one out. National survey information, suitably altered for known gross differences between the specific site and the nation as a whole -- such as behavioral differences related to climate -- will often give useful guidance. Other local data may be readily obtained on population distribution, economic level, types of industrial or other work activities in the area, geographic details, possible future developments, and so forth from such sources as local chambers of commerce, planning commissions, departments of motor vehicles, police departments, sheriff's offices or other local bodies, without the need for a physical visit to the site area; such information would provide further basis for the modification of nationally observed trends. Such information would also be helpful in deciding if a local survey was needed or not. Some regional and national survey data are under development. These data could be used to derive default exposure values in the absence of site-specific data, although the collection of site-specific data is encouraged in instances where site-specific conditions are unique and there is reasonable possibility that use of regional or national data could lead to large errors.

3.2 Chemical Mixtures

3.2.1 Additivity of Risk--Issue Six

An initial issue to address in considering additivity of risk is the accuracy and precision of the data on which the risk estimates are based. If the risk estimates are based on outdated, imprecise, or inaccurate studies, then their reliability will be constrained accordingly. Rigorous analysis of the existing scientific database requires not only diligence and time, but also adequate resources. The ultimate commitment for a Superfund site warrants a system that reflects state-of-the-art science and risk estimation.

The EPA Hazard Index (HI) depends upon Reference Doses (RfD) which, in turn, depend upon effect levels, such as the No Observed Adverse Effects Level (NOAEL), divided by an uncertainty factor. The ambiguities inherent in such a process had earlier led the Environmental Health Committee to urge greater reliance on the total dose-response function (when available) to calculate values such as Benchmark Doses (EPA, 1990; Crump, 1984) or effect level specifications that incorporate dose-response information (Barnes and Dourson, 1988).

A Hazard Index is often presumed to be a measure of the toxic potency of a mixture. But the RfD is only an indirect index of potency. It is only indirect because potency should be defined by the form of the dose-response (effect) relationship, not simply by a single, often equivocal, point on the dose-response function.

Given the interpretive flaws of the RfD, it is more appropriate to apply the strategy of the Benchmark Dose to mixtures. If a data set can support the derivation of a NOAEL, it should also provide enough information to calculate an ED_{10} or ED_{01}. (Comments by Dr. Gaylor discussed at previous Committee meetings indicate that, in fact, NOAELs approximate the lower confidence limits of the ED_{01})

Of course, exposure to any of the components of most environmental mixtures rarely approach even an ED_{01}, and the shape of the low-dose portion of the dose-response function is typically unknown. Moreover, the Hazard Quotient is not, as posed in Question 6, a risk estimate, and it is not risks that are added.

By utilizing dose-response information, more quantitative approaches would be possible. Such approaches might also help address the question of how to apply effect-specific RfDs. If separate dose-response functions can be fitted to individual toxic criteria, it should be possible to combine the available data in some form of meta-analysis. Another potential virtue of using dose-response information for individual effects is the possibility of examining potency ratios between effects at different positions of the dose-response function. Such an examination might offer clues about where to search for interactions. The Committee encourages EPA to utilize alternatives to the default approaches that involve additivity of risks or doses in specific cases whenever there is a reasonable scientific basis for so doing.

3.2.2 Hazard Indexes and Thresholds of Concern--Issue Seven

The Hazard Index approach to evaluating the non-cancer hazard of chemical mixtures is crude and affords far from universal protection for all mixtures. The Committee recommends, therefore, that if toxicological data on a particular mixture are adequate to derive an RfD for that mixture, the RfD should be used in place of the HI approach. Similarly, if a situation were to arise in which the scientific understanding of the interactions of chemicals comprising a mixture were sufficient to predict, theoretically, an RfD for a chemical based on its potential for interaction, the application of that RfD would be encouraged by the Committee.

However, the Committee also recognizes that neither of these cases is apt to occur frequently in practice. Therefore, some less precise approach, such as the use of the HI, seems unavoidable. The HI approach is not necessarily arbitrary, and in fact provides a very rational answer for an important class of interactions. Some of the useful features of the HI approach may be summed up in the three following characteristics. The validity of these characteristics is demonstrated in Appendix 1.

a) Characteristic 1: If all the chemicals in a mixture act toxicologically as if they are dilutions of a single chemical (This type of interaction will be referred to as "dilution-type" interaction), then the criterion HI = 1 should afford a level of protection that is intermediate between the levels of protection that are experienced as a result of exposure (limited to the levels associated with their respective RfDs) to the individual chemicals in the mixture. Consequently, if each RfD affords adequate protection for exposure to that individual chemical — as is intended in setting the RfD — then the condition HI = 1 should likewise afford adequate protection for exposure to the mixture.

b) Characteristic 2: If 1) each chemical in a mixture has an effective threshold when administered in isolation and 2) the RfD for each chemical in the mixture is below the threshold for that chemical and 3) the interactions among chemicals in the mixture involve a combination of independent mechanisms of action and dilution-type interactions (e.g., no synergistic interactions), then the threshold for the mixture should not be exceeded as long as HI $\leq$ 1. [See the Appendix for a rigorous definition of this class of chemicals.]

c) Characteristic 3: Even if all of the other conditions on the mixture in Characteristic 2 hold, if the hazard index of the mixture is greater than one (HI > 1), the threshold of the mixture may be exceeded.

Stated more broadly, Characteristic 2 implies that whenever the interactions in a mixture involve some combination of independent action and dilution-type interactions, then the HI approach will always afford at least as much protection as the least protective individual RfD. Characteristic 3 implies that under these types of interactions, the condition HI = 1 defines the least conservative approach that is still guaranteed to be protective, and that can be obtained without more detailed mechanistic information. (A less conservative approach [i.e., one consistent with HI > 1] can in some cases afford less protection than that afforded by any of the individual RfDs.)

An example of additional mechanistic information would be knowledge that all of the chemicals in the mixture operate through independent mechanisms. If that were known to be the case, then simply requiring the doses of the component chemicals in the mixture to be below their respective RfDs would be less conservative than the HI approach and still provide the same protection as that afforded by the individual RfDs. However, if we cannot rule out the possibility that some of the chemicals may operate by other mechanisms of joint action, then we cannot be sure that this approach affords adequate protection.

These considerations demonstrate that the value "1" (as in HI = 1) has a rational and meaningful basis for defining a "threshold of concern" for the HI that is not shared by any other number. Under the stated conditions, the uncertainty in the HI approach is, in a sense, no greater than that of the component RfDs. Given this, the Committee does not see any value, generally speaking, to use numbers other than "1" in defining a "threshold of concern." However, if there are interactions of the chemicals in the mixture which cannot be fully characterized by a combination of dilution-type interaction and independent mechanisms of action, then the entire HI approach may be inappropriate. Moreover, the use of the number "1" in defining a threshold of concern for a mixture, as derived in the Appendix, does not take into account the possibility that the joint severity to a particular individual subjected to a number of unrelated adverse effects, whether they exhibit dilution type interactions or not, may be very great, and the safety of the mixture may therefore be questionable.

It should also be noted that none of these properties assume a linear dose response. The hazard index approach is not predicated upon a linear dose response.

3.2.3 Interpretation of the Hazard Index--Issue Eight

When the hazard index (HI) is greater than unity, the toxicologic connotations and interpretations may be different depending on several factors:

a) If two or more agents involved share the same mechanism of toxicity their doses could well be additive, possibly resulting in more than the sum of the additive risks (see discussion in the Appendix).

b) If they instead act upon two or more sites along a mechanism sequence, they could well be supra-additive, as are Malathion/Parathion (Murphy, 1969).

c) If they each act by different toxicologic mechanisms, additivity of risks for a common endpoint is not necessarily to be expected; for different endpoints, the potential hazard could be overestimated.

It should be noted that the occurrence of agents acting in the same target organ does not imply that they share the equivalent mechanistic properties. For instance, there are many different types and mechanisms of hepatotoxicity. One should never consider that because two agents cause reproductive (e.g., testicular) or developmental toxicity (e.g., cleft palate), they operate by the same or even similar mechanisms. On the other hand, diverse types of developmental (for instance) toxicity can result from the same apparent mechanism, e.g., adenosine tri-phosphate deficiency and mitotic arrest, or interference with directional cell migration.

In the absence of experimental data sets, it is not yet possible to contemplate whether the slope of increasing risk would be steep or shallow. With adequate experimental data sets, one could still have different levels of concern, depending on both the slopes of the dose-response curves for mixture components as well as the types of resulting effects. The most prudent mode would be to consider that as the HI exceeds unity, the potential for risk increases; without a more complete understanding of interaction mechanisms, however, we can not state how rapidly this increase occurs. Also, we cannot rely on comparisons of risk using HIs for HIs greater than unity.

3.2.4 The RfD As a Criterion of Toxicity--Issue Nine

The OSWER proposal suggests that EPA modify its current practice of calculating the Hazard Index for mixtures. At present, chemicals in a mixture are assessed for joint action by computing the ratio of exposure to RfD, then adding these quotients to obtain the Hazard Index. Because the RfDs are based on doses such as NOAELs that are derived from the critical effect in an assay (that is, the effect showing the greatest sensitivity to exposure), the resulting Hazard Index may encompass a spectrum of toxic endpoints and risk levels. Such a melding of disparate endpoints may overestimate the magnitude of risk. The proposed modification would calculate

effect-specific RfDs instead, and is stated to be more consistent with the assumption of dose-additivity that guides the Hazard Index calculation.

The argument presented in the Issue paper has merit, but offers practical difficulties. One reason is that some RfDs are based on non-specific endpoints such as weight loss, which can arise from many causes. Another is the problem of effect interactions; for example, kidney damage may lead to peripheral neuropathy, and liver dysfunction may prolong the CNS action of a solvent.

More descriptive, flexible indices of joint action might prove useful. Assume, for example, at least 100 identifiable contaminants at a site. If each is present in groundwater, say at an average of 1% of its RfD, does it make sense to add all the Hazard Quotients to derive a single sum expressed as a Hazard Index of 1.0? Would it be more reasonable to attempt to derive some estimate of the degree to which these substances act jointly, or embrace common modes of toxicity?

As an example of such an approach, assume a mixture of five chemicals, each present at 0.2 of the individual RfDs. If the biological effects of all five were totally independent, that is, affecting different organs, systems, or receptors, then the total effect of the mixture would still be 0.2 times the RfD. If the biological effects of all five overlapped completely, and acted in the same way on the same system, then the total effect of the mixture would be the sum of the individual RfDs, or 1.0.

Another approach would be to conceive of the overlap estimates as correlations. Under this assumption, the matrix might then be subjected to Principal Components analysis. That is, the cell entries would represent shared variance.

Assume, however, that the mutual overlap = 0.01. That is, any one of the chemicals enhances the effect of one of the others by 1.0%. Then, the combined effects of chemicals 1 and 2 = 0.2 + 2 (.01) = 0.22, because chemical 1 enhances chemical 2 by 1.0% and the reverse.

Filling in each of the interaction cells in the 5 x 5 matrix (10 cells) yields .02 x 10 = 0.2. Therefore, the total proportion of biologically effective RfD is 0.4 rather than 1.0, which would be the current default Hazard Index.

Perhaps such a value could be expressed by a matrix (100 x 100 in this example) in which each cell contains an estimate of overlap or joint action, defined elsewhere as "commonality" (Weiss, 1986). We would expect commonality coefficients within chemical classes, for example, to exceed commonality estimates between classes. Although these estimates would be expert judgments, they need not be any more vague than the EPA approaches to probabalistic risk assessment or any less cogent than Office of Toxics Substances' reviews of chemical structures. Such a matrix might substitute for, or be an accessory to, effect-specific RfDs.

A first step in this ambitious project would be to group chemicals for specific effects into groups that are likely to operate through some combined (i.e., non-independent) action. The next step would be to assign some type of interaction term to all pairs of chemicals in the same group as suggested above. This step is likely to be more difficult as well as controversial. Until these interaction terms are developed and adopted, EPA could make an incremental improvement just by taking advantage of the presumption that chemicals in different groups operate by independent mechanisms of action, but otherwise retaining other facets of the current approach. The result would be an approach that is similar to that proposed by EPA, except that chemicals which cause the same endpoint would not necessarily be assumed to operate by a common mechanism of action.

Finally, the proposed modifications remain captives of NOAELs and their relatives. Contrary to what is stated in the issue paper, effect levels are not thresholds as is noted repeatedly in EPA documents. Examine the case study of chloroform -- the LOAEL for liver damage is 12.9 mg/kg daily, and the RfD is calculated on the basis of an uncertainty factor of 1000. The NOAEL for kidney effects (based on the same dog study) is also 12.9 mg/kg, but in accordance with EPA practice, it is divided by an uncertainty factor of 100. The comparisons imply the assignment of a severity index that probably is unjustified on the basis of these kinds of data; moreover, the basis for such an assignment is not adequately explained. The absence of dose-response factors is a further defect.

Despite these problems with the proposed approach, it is not nearly as seriously flawed as the two alternatives presently in use. The "critical" effect approach ignores the fact that a chemical can cause an effect even when that effect is not the critical effect and therefore is not health protective. On the other hand, the "Superfund" approach assigns the RfD of the critical effect to all effects of a chemical, which

is clearly overly conservative. The proposed approach, which involves calculating an RfD for each organ site for each chemical and then combining over the chemicals for each organ site, and then combining over organ sites is, despite its flaws, an improvement over either of these two alternatives.

RfDs for individual agents have clear value in risk management in setting "safe" limits of exposure for such agents (that is: very low, essentially zero -- or perhaps even zero -- risk limits) and HQs for individual agents clearly measure how far from or close to the RfDs the actual exposures are for those agents. The HQs give a measure of risks at the actual exposures relative to those if the exposures were equal to the RfDs -- at least for HQs above 1.0 and possibly below, depending on the existence or nonexistence of thresholds and where the RfDs are located with respect to the thresholds. Ideally, in assessing a mixture of agents using the RfD concept, the RfD of the mixture, as determined from exposures using the mixture itself, is what is needed to assess the safety or lack of safety (the lack or presence of risk) of actual exposures to the mixture. If such a "mixture RfD" were available, then the HQ for the mixture would indicate risk relative to the RfD, an HQ of 1.0 being, again, the dividing line between "safe" and "unsafe". Experimentally determined RfDs for specific mixtures are generally not available, and it is not generally practical during a particular assessment to obtain them. It has been suggested that HIs greater than 1.0 should suggest particular concern, an HI for a mixture, in this sense, being taken as a kind of surrogate for the actual HQ of the mixture.

Considering two cases helps answer the questions raised here: (a) in deriving RfDs from experimental data the original idea was that, for non-carcinogenic adverse effects, thresholds exist and that the RfDs represent safe doses located somewhere below the threshold doses for each agent -- how far below not being known; and (b) it may be, instead, that there is residual risk below the RfD because there is no threshold or, alternatively, that there is a threshold but the RfD is set, unintentionally, somewhere above it.

In case (a), it is easy to show, with a few numerical examples, that a higher HI does not necessarily imply a higher risk. Considering (b), the actual risk -- the probability of a particular adverse effect -- corresponding to the RfD is not known and can differ significantly from one RfD to another.

Nothing in the various procedures for determining RfDs is aimed at ensuring that RfDs are determined at either known levels, or a fixed level, of risk. Comparisons of HQs for different agents is therefore not meaningful from a risk standpoint since the ratios depend on the RfDs determined at varying but unknown levels of risk and reflect this variation. An HQ of 3.0 for one agent, for example, might not indicate a higher risk than for another agent having an HQ of 2.0 if the RfD for the first agent was, by chance, determined at a sufficiently lower value of risk than was the second (depending, also, on the shapes of the dose response curves at low doses). Summing the HQs to produce an HI inherently assumes that the HQs are comparable. Since it cannot be generally known that there is comparability of HQs and since comparability does not therefore exist in general, the significance of the sums is, in turn, in doubt. It cannot be known whether a given mixture having a larger HI than another mixture represents a greater or lesser risk compared to that mixture. If it were possible to derive RfDs at some standardized risk level, it might be possible to develop a basis for comparing HQs and for forming meaningful HIs. In the case that the dose responses of individual mixture components at low doses are linear and extend to zero, for example, HQs are just the ratios of the probabilities of effect at the actual exposures to those at the RfDs; if the RfDs were all determined at a standardized probability level then the HIs would be proportional to the sum of the probabilities of effect at the actual exposures and under these circumstances the HIs would offer measures of relative risk. No accounting for synergism, antagonism or of joint severity is included in this approach (nor are they in the HI as now defined), and the implicit assumption of dilution-type interaction may well be invalid for many interactions.

From the foregoing discussion, we can state that HIs do not have general meaning with regard to relative risk, and their undiscriminating use can lead to giving a degree of unwarranted comfort to the unwary. A high HI may or may not mean a high relative risk and a low one must not be taken to mean that the relative risk is necessarily low. The contributions of individual agents need to be considered individually; in addition, the combined effects of agents present need also to be considered, taking into account at least the modes of action, when known, of the organs or systems affected, and the possible joint severity of the effects -- all this aside from any synergism or antagonisms that may be present. Even if synergism and antagonisms are not present, the total severity of being affected independently by more than one adverse effect (or by effects caused by more than one agent) can be much greater than the severities of the individual effects might lead one to believe; and the risk (defined as including both severity and probability) will therefore be much

greater, too, than might be anticipated by considering the array of individual HQs, alone.

Mixtures containing one or more agents, the HQs of which are greater than 1.0 may pose a risk on the basis of the RfD concept; those not containing such agents cannot be considered risk-free, however, and must be subjected to further examination -- as must the former, since reducing exposures to reduce the HQs greater than 1.0 to values less than 1.0 will not necessarily bring the total risk to a low enough level.

Under conditions described in full in the Appendix and elsewhere in the body of this report, a Hazard Index of 1 for a mixture can have special meanings in relation to the RfDs of the individual chemicals, the criterion HI = 1 affording a level of protection that is at least as protective as the least protective RfD for a single chemical in the mixture. However, this may not be the case if the HI exceeds 1.0 for a mixture. Moreover, if some of the chemicals interact synergistically, then the condition HI = 1.0 may not afford adequate protection. On the other hand, if the individual chemicals have independent mechanisms of action, then the criterion HI = 1.0 may be overly protective. Whenever there is no, or very limited, information on the types of interactions that exist among the chemicals within a specific mixture, the Committee suggests that the HI, including the criterion of HI = 1, be considered (along with whatever other information may be available) for possible guidance unless information becomes available which confirms the validity of HI = 1, or unless information on interactions becomes available which permits the development of an appropriate value of HI or another criterion.

The Committee realizes that most Superfund decisions must be made on the basis of incomplete (and usually highly imperfect) information. This situation is, of course, the root cause of the estimation problem with which we must cope.

3.3 RfDs in Goal Setting

3.3.1 Exposure Scenarios--Issue Ten

OSWER initially considered two approaches for using RfDs in setting risk-based remediation goals in soil: 1) comparison of a 6-year, childhood exposure to contaminants in soil with a sub-chronic RfD; and, 2) comparison of a 30-year, time-weighted average exposure to contaminants in soil (including exposures to both children and

adults) with a chronic RfD. Now, a third approach has been proposed: comparison of a 6-year, childhood exposure with a chronic RfD.

The second approach proposed by the OSWER probably is the more reasonable. That is, to compare a 30-year time-weighted average (TWA) exposure with a chronic RfD. It is likely to be adequately conservative. Comparison of a six-year old's exposure with a chronic RfD (a third approach) may be overly conservative. It also assumes the six-year old is the more vulnerable. The second approach accounts for variable susceptibility with age in a more conservative manner than does method three. Actually, all three methods could be considered, and the one giving the most conservative (in the absence of specific information) or the most reasonable estimate (in the presence of such information) used. It would be helpful to see a group of diverse examples for all three approaches. Perhaps the most relevant exposure scenario could then be selected (i.e., childhood vs. lifetime). Clearly, the model selected must be one that accommodates the most intense future land utilization, e.g., housing, lest repeat remediation become necessary.

3.4 Short-Term Toxicity Values

3.4.1 Interim Estimates of Toxicity--Issue Eleven

The methods outlined in the *Interim Methods for Development of Inhalation Reference Concentrations* are a reasonable approach to determining short-term toxicity values. Basically, the approach is to find human toxicity data (if possible) or animal data of the appropriate duration (or as close as possible to the appropriate duration) that indicate a NOAEL. The NOAEL is then used to set the RfC, based on the NOAEL (adjusted for duration) divided by uncertainty or modifying factors. Care must be taken in adjusting the NOAEL for the duration of exposure for compounds that cause acute effects based mainly on concentration and not duration of exposure. As noted in section 3.3.1, concerning time-weighted averaging, caution must be used when the relevant extrapolation ranges are fairly broad.

A major concern in the calculation of short-term toxicity values is to choose appropriate uncertainty and modifying factors so as not to exaggerate the potential toxicity associated with a site or a specific chemical and yet protect the public health. It will also be important in assessing risk from short-term exposures that the exposure

time and the averaging time used are consistent with the shortest period of time that could produce an effect (section 6.4.2 of RAGS-A).

There are two additional sources of information on short-term toxicity values. One is the Emergency Response Planning Guidelines (ERPG) of the American Conference of Governmental Industrial Hygienists. The second is the Community Emergency Exposure Levels (CEELs) being developed by a National Academy of Sciences Committee sponsored by the EPA. The approaches of both of these committees should be considered in developing short-term toxicity values. Although some of the short-term exposures of interest to the EPA at Superfund sites are of a longer duration than one to eight hours, the concept of providing some measure of the dose-response characteristics of a chemical is an excellent one. For persons in risk management, it is important to know if the level of a chemical that causes mild irritation is two times or one thousand times lower than the level that is life threatening. The ERPG and CEEL methods use multiple guidance levels (ERPG-1, ERPG-2, ERPG-3 or CEEL-1, CEEL-2, CEEL-3) to provide some dose-response information for the chemical of concern. The lower guidance level, such as CEEL-1 or ERPG-1 is the level below which the chemical is unlikely to cause mild effects such as discomfort or irritation. The second level is the level below which the chemical is unlikely to cause toxic effects leading to disability that could interfere with taking protective actions. The third level is the level that is life threatening. These short-term toxicity guidance levels are set for specific times of exposure (one or eight hours).

3.4.2 Short-Term Air Action Levels--Issue Twelve

The method proposed by Region 6 suggests the possible use of OSHA standards, such as Permissible Exposure Limits (PELs) and Short-Term Exposure Levels (STELs), to derive short-term action levels when alternative EPA RfD data are not available. Clearly the absence of an RfD does not mean that there should be no short-term action level. Ideally the data used to derive the OSHA standard should be examined to see if the same data can be used to derive an appropriate RfD. The use of EPA derivation methods would help promote consistency across various hazardous substances; hence the use of EPA methods should be encouraged. Where this is not possible or practical, the use of health data on which OSHA standards have been based could be considered, taking into account differences between the worker population and the general population. In its consideration of adapting EEGLs to the general population, the National Research Council (*Criteria and Methods for Preparing*

Emergency Exposure Guidance Level (EEGL), etc., 1986) applied a safety factor of two for sensitive subgroups of the general public and a safety factor of ten for newborn infants. The rationale used to derive these safety factors should be reviewed to see if it is appropriate for PELs and STELs as well.

4. CONCLUSIONS

The OSWER's attempts to codify, and so increase the consistency, of its site-specific risk assessment methodology are praiseworthy. The need to deal with wide ranges of contaminants, often in complex mixtures with exposure through multiple media, poses daunting problems for risk assessment. Although the OSWER has (understandably) not succeeded fully in meeting the goals they set for themselves in producing the HHEM document, they have made a good start. The following comments summarize the Committee's major findings, specifically noting where we believe a redirection in approach is called for, and providing advice where possible.

The Committee is of the opinion that there are some serious difficulties, both conceptual and practical, with the approach recommended in the RAGs document for calculating an RME based on a UCL on the average concentration at a site.

First, a UCL for the mean concentration does not lead logically to a "reasonable maximal exposure." The current approach assumes that the samples taken are representative of those areas where exposures are most likely to occur. The RME is also a function of the number of samples available; a larger number of samples will result in a smaller RME even if the samples are not representative of exposure opportunities. For example, if a site is well-characterized, so that the UCL is very close to the true mean, but a "hot spot" is very attractive to visitors, then a significant fraction of visitors to the site (perhaps the majority) could be exposed to levels significantly higher than the UCL.

Second, to calculate the RME, the UCL on concentration is combined with 90th percentile values for some variables and 50th percentile values for other variables in an *ad hoc* fashion, making the resulting RME very difficult to interpret.

Third, the UCL on the mean concentration does not take into account the spatial distribution of contamination *via-a-vis* the relative frequency with which people are likely to visit various parts of the site.

Fourth, the calculation of a UCL requires statistical assumptions that are generally not met by sampling plans at a Superfund site. As a consequence, statistical procedures can sometimes produce a UCL that has little, if any, relation to actual concentrations at a site.

Because of these difficulties, the Committee recommends that the EPA move towards a "full distributional approach" to calculating the RME. In such an approach, a distribution is developed for each of the input terms in the exposure equation. These distributions determine a distribution for exposure, which can be calculated using Monte Carlo methods. A particular percentile of this exposure distribution, such as the 90th percentile, could be used as the definition of the RME.

If people are more likely to visit certain areas of the site than others, then the spatial distribution of contaminant concentrations about the site needs to be considered when quantifying human exposure. Kriging and triangulation are two statistical methods for quantifying concentration that take into account the spatial distribution of samples. These approaches should be considered for adoption by EPA.

As part of the effort to implement a distributional approach for quantifying exposure, EPA should develop default distributions for exposure parameters that are unlikely to vary significantly from site to site. However, the collection of site-specific data is encouraged in instances where site-specific conditions are unique and there is a reasonable possibility that use of default distributions developed from regional or national data could lead to gross errors.

The Committee recognizes that some time will be required to implement a distributional approach to quantifying exposure. In the meantime, some version of the current approach may have to be used. The Committee has two recommendations regarding the application of the current approach during this interim period. First, as long as some type of mean concentration is to be employed to estimate human exposure, the Committee agrees with OSWER that an arithmetic mean is more appropriate than a geometric mean. Second, given the difficulty in interpreting the RME as presently calculated, the Committee recommends that some type of 'most reasonable' estimate of exposure also be calculated and made available to risk managers along with the RME.

The issue of risk additivity from exposure to complex mixtures remains a difficult question. The Committee is concerned about the approach of using RfD-derived Hazard Quotients/Hazard Indices as a basis for adding "risks." In our opinion, the Hazard Quotient is not a risk estimate, and it is not truly risks which are being added when the proposed approach is used. We would much prefer to see quantitative applications using dose-response data (not the "point" data represented by

LOAEL/NOAEL-derived RfDs) to drive risk estimates. Further, we suggest that the Agency develop and use alternatives to the current default approaches that assume risk additivity whenever there is a reasonable scientific basis for so doing.

The use of the HI itself can be misleading, and it should used as a "fallback," with full recognition of its possible inapplicability, only when more refined toxicological data are not available. As noted in section 3..2.2 of the report (and demonstrated in the Appendix), the condition "HI = 1" defines a "threshold of concern" that is not shared by any other value for HI, and for which, under specified conditions, the uncertainty in the HI approach is no greater than that of the component RfDs. The HI approach is invalid, however, if the chemicals in the mixture cannot be fully characterized by a combination of dilution-type interactions and independent mechanisms of action. Another instance in which the HI = 1 approach may not adequately address the safety of a mixture occurs when there is a high joint severity of total effect in individuals subjected to a number of unrelated adverse effects (regardless of the mechanisms of action).

Interpretation of an HI greater than 1 can vary depending on several toxicological factors: the existence of common mechanisms of toxicity (in which case doses could be additive; cases in which the agents act upon two or more sites along a mechanism sequence (resulting in supra-additivity); and instances in which the substances act by different toxicologic mechanisms (in which case, potential hazard could be overestimated. Although it is likely that risk increases as the HI exceeds 1, we can not state without a more complete understanding of interaction mechanisms) how rapidly this increase occurs. For this same reason we cannot rely on HI-based comparisons of risks for HIs greater than 1.

The proposed guidance recommends modifying current policy and using the RfD as the toxicity criterion for each of the other effects believed to be caused by a given agent in a chemical mixture. This proposal is not without merit, since the current approach melds many disparate endpoints and may overestimate risk. The proposal suffers however, from problems of its own--it does not deal with effect interactions, nor with the fact that many RfDs are based on non-specific endpoints which can stem from many different causes. The proposed approach, despite its limitations, is an improvement over the alternatives presently in use.

OSWER discusses three approaches for using RfDs to develop risk-based remediation goals for contaminated soil, involving differing exposure scenarios and target populations: 1) comparison of a 6-year, childhood exposure to contaminants in soil with a sub-chronic RfD; 2) comparison of a 30-year, time-weighted average exposure to contaminants in soil (including exposures to both children and adults) with a chronic RfD; and 3)comparison of a 6-year, childhood exposure with a chronic RfD.

The most reasonable and supportable approach appears to using a 30-year time-weighted average with a chronic RfD, but differences between the three proposals are not dramatic. OSWER should study all three approaches, applied to a diverse set of examples in order to verify its ultimate choice (or range of choices).

The Committee sees no particular problems in the existing approach for dealing with short-term toxicity estimates. As in all cases when using time-weighted averaging, care must be taken when the extrapolation ranges are broad. The Committee suggests that OSWER take cognizance of the EPA-sponsored work at the National Academy of Science on Community Emergency Exposure Levels, and of the work on Emergency Response Planning Guidelines by the American Conference of Governmental Industrial Hygienists.

The method proposed by Region 6 for setting short-term air action levels calls for the possible use of OSHA standards, such as Permissible Exposure Limits (PELs) and Short-Term Exposure Levels (STELs). The data used to derive any OSHA standard so used should be examined to see if the same data can be used to derive an appropriate RfD. The use of EPA derivation methods would help promote consistency across various hazardous substances; hence the use of EPA methods should be encouraged. Where this is not possible or practical, the use of health data on which OSHA standards have been based could be considered, taking into account differences between the worker population and the general population.

6. REFERENCES

Barnes, D.G., and M. Dourson. 1986. Reference Dose (RfD): Description and use in health risk assessments. *Regulatory Toxicol. Pharamacol.* (8): 471-486.

Crump, K. S. 1978. Estimation of Mean Pesticide Concentrations When Observations are Detected Below the Concentration Limit. Prepared for the Food and Drug Administration.

Crump, K.S. 1984. A new methods for determining allowable daily intakes. *Fund. Appl. Toxicol.* (4):854-871.

EPA (Environmental Health Committee, U.S. EPA Science Advisory Board). 1990. *Use of Uncertainty and Modifying Factors in Establishing Reference Dose Levels.* EPA-SAB-EHC-90-005.

Murphy, S.D. 1969. Mechanisms of pesticides interaction in vertebrates. *Residue Reviews.* (25):201-221.

National Research Council. 1986. Criteria and Methods for Preparing Emergency Exposure Guidance Levels (EEGL).

Weiss, B. 1986. In: Annau, Z. (ed.) *Neurobehaviorial Toxicology.* Baltimore: Johns Hopkins University Press. pp. 1-20.

APPENDIX A - CHARACTERISTICS OF THE HAZARD INDEX

If n chemicals in a mixture operate by "dilution-type interaction" (i.e., have the same toxic properties as if they were all dilutions of the same chemical), then the probability of an adverse effect from simultaneous exposure to doses E_1, ..., E_n of the n respective chemicals can be expressed in terms of a single dose response, P(d),

$$P(E_1T_1+\ldots+E_nT_n), \tag{1}$$

where T_1, ..., T_n are toxicity equivalent factors (TEFs) for the respective chemicals. TEFs are used by EPA to assess risks of mixtures such as those comprised of dioxin, furans and some PCBs that are thought to have a common mechanism of action. Note that this formulation makes no assumptions regarding the shape of the dose response, P(d). E.g., it can be linear, nonlinear or even include a threshold. The shapes of the dose-response curves of individual components, relative to each other, must be such as to result in the "dilution type interaction."

It is clear from this formulation that the RfD for chemical i is equally protective as the RfD for chemical 1 if

$$R_iT_i=R_1T_1 \tag{2}$$

where R_i is the RfD of chemical i (that is, that each chemical, at its RfD, elicits the same response as others do at their RfDs).

The validity of Characteristic 1 is now demonstrated. Suppose that a mixture has HI = 1 and that all RfDs from each of the n chemicals are equally protective. That means that the component doses E_i satisfy

$$HI = \frac{E_1}{R_1} + \ldots + \frac{E_n}{R_n} = 1 \tag{3}$$

If we solve for R_1 in expression (1) for i > 1 and substitute this for R_i into expression (2), we get

$$HI = \frac{E_1}{R_1} + \frac{E_2T_2}{R_1T_1} + \ldots + \frac{E_nT_n}{R_1T_1} = 1 \tag{4}$$

or, equivalently,

$$E_1T_1 + E_2T_2 + \ldots + E_nT_n = R_1T_1 \qquad (5))$$

However, this expression says that exposure to the mixture provides exactly the same risk as exposure to the RfD of chemical 1 alone, which is by assumption the same as the risk from exposure to the RfD of any other of the n chemicals.

Next consider the case in which all RfDs of the n chemicals are not equally protective (i.e., the product of the RfD and the corresponding TEF are not the same for all chemicals). Suppose, without loss of generality, that $R_1T_1 \leq R_iT_i$, or, equivalently, $R_i \geq R_1T_1/T_i$, for i=1,...,n. Replacing R_i, i>1, in equation (2) by the right side of this inequality yields

$$\frac{E_1}{R_1} + \frac{E_2T_2}{R_1T_1} + \ldots + \frac{E_nT_n}{R_1T_1} \geq 1 \qquad (6)$$

or, equivalently,

$$E_1T_1 + \ldots + E_nT_n \geq R_1T_1 = MIN(R_1T_1, \ldots, R_nT_n) \qquad (7)$$

Similarly,

$$E_1T_1 + \ldots + E_nT_n \leq MAX(R_1T_1, \ldots, R_nT_n \qquad (8)$$

These two inequalities indicate that exposure to the mixture with an HI of 1 poses a risk that is intermediate between that posed by the most protective and the least protective of the RfDs of the individual chemicals. However, if the RfD for each individual chemical is sufficiently protective, then the protection afforded by the condition HI = 1 must be sufficiently protective as well (because it is more protective than the least protective RfD).

Next we demonstrate the validity of Characteristic 2. The assumption here is that the chemicals in the mixture can be divided into subgroups, such that chemicals within each subgroup have dilution-type interactions and chemicals in different subgroups have independent mechanisms of action. Independence of action implies that thresholds for a given chemical are unaffected by exposures from other chemicals having independent mechanisms of action. I.e., if a given dose of a particular chemical (or combined dose of chemicals having dilution-type interaction) is below the threshold for an effect (i.e., does not increase the likelihood of an effect) when given in

isolation, then that same dose will have no effect when given in combination with other chemicals having independent mechanisms of action.

Now suppose a mixture has HI = 1 and is composed of doses E_1, E_2, ..., E_n of its component chemicals. This means that the component doses satisfy expression (2). Without loss of generality, suppose that E_1, E_2,...,E_m, where $m < n$, are the component doses from one of the subgroups of chemicals that have dilution-type interactions. Since HI = 1 for the complete mixture, it follows that, for this subgroup, we have

$$\frac{E_1}{R_1} + \ldots + \frac{E_m}{R_m} \leq 1 \qquad (9)$$

which, reasoning exactly as before, implies that

$$E_1T_1 + E_2T_2 + \ldots + E_mT_m \leq \text{MAX}(R_1T_1, \ldots, R_mT_m) \qquad (10)$$

This, in turn, implies that exposure to doses E_1, ..., E_m in this subgroup is no less protective (i.e., is no more likely to cause an adverse effect) than exposure to the least protective RfD in this subgroup. But if all of the RfDs are below the respective thresholds, then exposure to this sub-mixture will likewise not increase the likelihood of an adverse effect, at least when exposure to the sub-mixture is in the absence of other exposures. However, since we are assuming the other chemicals have mechanism of action that are independent of those in this subgroup, exposure to this sub-mixture will have no effect even in the presence of the remaining chemicals in the mixture. This demonstrates the validity of Characteristic 2.

To demonstrate the validity of Characteristic 3, we simply note that if a mixture is comprised of chemicals with dilution-type interactions and if the RfD for each chemical happens to coincide with this threshold, then, by the same reasoning that led to expression (4), we have

$$E_1T_1 + E_2T_2 + \ldots E_nT_n > R_1T_1 \qquad (11)$$

which implies that the threshold for the mixture has been exceeded.

APPENDIX B - GLOSSARY

CEEL	Community Emergency Exposure Levels
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
CNS	Central Nervous System
ECAO	Environmental Criteria and Assessment Office
ED_{01}	Effective Dose (level at which 1% of test animal subjects are affected)
ED_{10}	Effective Dose (level at which 10% of test animal subjects are affected)
EEGL	Emergency Exposure Guidance Level
EPA	Environmental Protection Agency
ERPG	Emergency Response Planning Guidelines
HHEM	Human Health Evaluation Manual
HI	Hazard Index
HQ	Hazard Quotient
LOAEL	Lowest Observed Adverse Effects Level
MEI	Maximally Exposed Individual
MOV	Maximum Observed Value
NOAEL	No Observed Adverse Effects Level
OSHA	Occupational Safety and Health Administration
OSWER	Office of Solid Waste and Emergency Response
PEL	Permissable Exposure Limit
PRP	Potentially Responsible Parties
RAGS	Risk Assessment Guidelines for Superfund
RfC	Reference Concentration
RfD	Reference Dose
RME	Reasonable Maximum Exposure
STEL	Short-Term Exposure Levels
TEF	Toxicity Equivalent Factor
TWA	Time-weighted Average
UCL	Upper Confidence Limit

Distribution List

Administrator
Deputy Administrator
Assistant Administrators
Deputy Assistant Administrator for Research and Development
Deputy Assistant Administrator for Water
EPA Regional Administrators
EPA Laboratory Directors
EPA Headquarters Library
EPA Regional Libraries
EPA Laboratory Libraries